

Lo spazio cibernetico come nuovo dominio bellico

The dark side of CyberWorld



Gian Piero Siroli, Physics & Astronomy Dept. Univ. of Bologna & CERN

USPID - Dicembre 2021

What a cyber-weapon can look like: Stuxnet

- A “worm” designed to sabotage a specific industrial process. It penetrates a particular subsystem of a SCADA industrial control systems of a single producer (Siemens). Once injected, it spreads silently in the Windows/SCADA infrastructure looking for specific Programmable Logic Controllers (PLC) and reprogram them to alter the functionality, showing at the same time normal running conditions to the monitoring system

Currently no formal international consensus on definition of “cyber-weapon”, but for any practical purpose “computer code used, or designed to be used, with the aim of threatening or causing physical, functional, or mental harm to structures, systems, or living beings”

Cyber-Weapons” T.Rid & P.McBurney 2012

- Disruption of Iran's nuclear program by damaging centrifuges at uranium enrichment facility in Natanz
- Worm analyzed in public conferences, papers from various authors, probably the best studied piece of malware in history. Executable code available on the network

...and many more...

Significant cyber incidents since 2006 (government agencies, defense & high tech companies, cyber-crime >1M\$ loss)

(Center for Strategic & International Studies CSIS)

October 2021. A Chinese-linked hacking group gained access to calling records and text messages from telecommunication carriers across the globe, according to a report from CrowdStrike. The report outlines the group began its cyberattacks in 2016 and infiltrated at least 13 telecommunications networks.

October 2021. A cyberattack targeted the government-issued electronic cards Iranians use to buy subsidized fuel and altered the text of electronic billboards to display anti-regime messages against the Supreme Leader Ayatollah Ali Khamenei.

October 2021. A group with ties to Iran attempted to hack over 250 Office365 accounts. All the targeted accounts were either U.S. and Israeli defense technology companies, had a focus on Persian Gulf ports of entry, or maritime transportation companies with a presence in the Middle East.

October 2021. Brazilian hackers carried out a cyberattack on the National Malware Center website belonging to Indonesia's State Cyber and Password Agency. The hackers edited the contents of the webpage and indicated that the cyberattack was retribution for an Indonesian hack on the Brazilian state website.

October 2021. Hackers leaked data and photos from the Israeli Defense Ministry after gaining access to 165 servers and 254 websites, overall compiling around 11 terabytes of data.

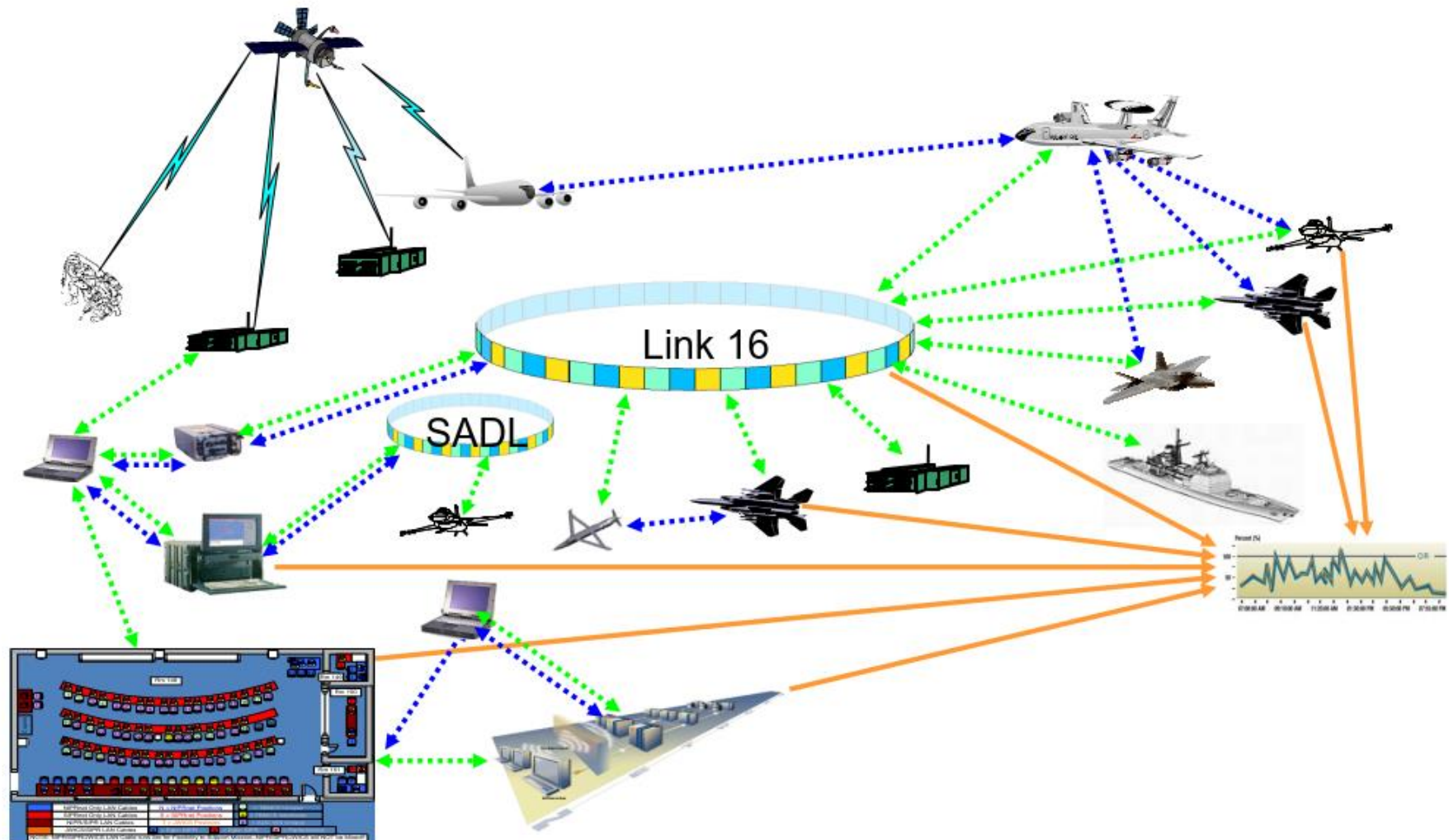
....



Full Battlespace Test

Battlefield digitization (USAF 2017)

Command & Control, Datalinks, Weapons, and Interactions







Battlefield digitization

Project to develop large-scale battle management software to enable military planning at the theater level

- **Jan 2021: DARPA awarded a \$10.4 million contract to Raytheon Corp. for Joint All-Domain Warfighting Software (JAWS) program**
 - **Dynamic coordination of kill webs across battlespace (sensors, weapons, & decision makers are not always in the same place)**
 - **Developing s/w to set up synchronized kill webs operating on and under the sea, on land, in the air, in space, and in the electromagnetic spectrum**
 - **Primary issues for C&C: sensing, communications, weapons**
 - Sensing is the ability to detect, geolocate, and identify potential targets for attack
 - Communications ability to pass data at the right time and quickly enough to deploy, maintain, and maneuver assets
 - Weapons involves choosing right weapon for the task
 - Restore system operations, increase resilience, regain initiative. Difficulty of attribution, attacks from third party systems
 - **JAWS seeks to develop s/w tools to create a distributed C&C structure using dynamic teaming and machine-to-machine interfaces to enable centralized and distributed planning and execution**

Cyber-(in)security of nuclear weapons

Key cyber vulnerabilities and potential consequences

	POINT OF VULNERABILITY	TYPE OF ATTACK	POTENTIAL CONSEQUENCE	
	Early Warning Systems: Radars and Satellites	Spoof of an incoming nuclear attack	Nuclear launch based on false warning	
	Communications Systems	Cyberattack disrupts or disables communication channels between officials, operators/systems, international counterparts	Nuclear launch based on misinterpretation of information/inability to de-escalate crisis OR Loss of confidence in ability to issue launch orders to respond to nuclear attack	
	Supply Chain	Malware or malicious code introduced into a nuclear weapon component	Loss of confidence in nuclear weapon operating as intended	
	Security Systems	Cyberattack disables or defeats physical security measures	Theft of nuclear weapon	

“Nuclear weapons in the new cyber age”

Report of the cyber-nuclear weapons study group (NTI)

Autonomous Weapon Systems (AWS)

- **Lethal Autonomous Weapons Systems (LAWS):** a system that, once launched/activated, “can select targets and apply force without meaningful human control” (International Committee for Robot Arms Control)

“Lethal autonomous weapons threaten to become the third revolution in warfare. Once developed, they will permit armed conflict to be fought at a scale greater than ever, and at timescales faster than humans can comprehend. These weapons can be used against innocent populations and hacked to behave in undesirable ways. Not much time to act, once this Pandora’s box is opened, it will be hard to close.

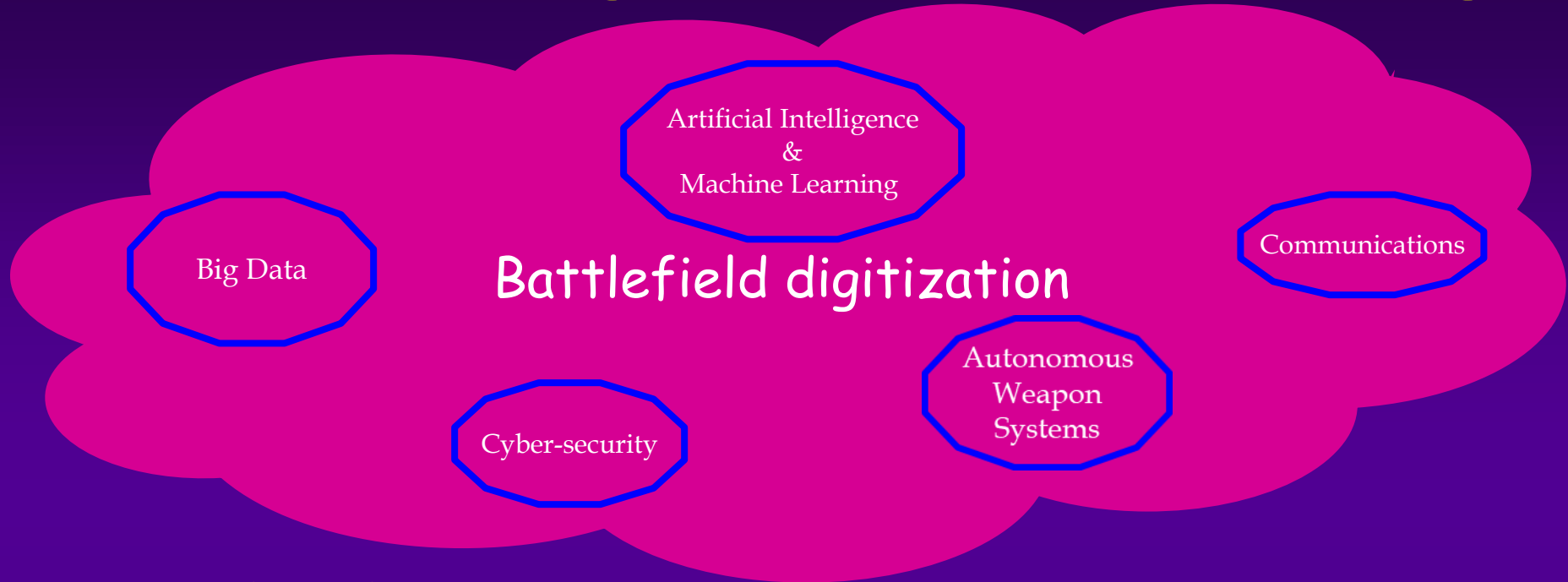
International humanitarian law continues to apply fully to all weapons systems, including the potential development and use of lethal autonomous weapons systems”

REMINDER

In the cyber domain full autonomy is already operational

- **Open Letter to the UN Convention on Certain Conventional Weapons (CCW) by tech companies on AI & robotics**
Attacks can be delivered at a non-human time scale and the speed in which one is able to respond is of equal importance as time to detect

Artificial Intelligence – Machine Learning



- **AI - defense of critical Networks: real time, pattern recognition, anomaly detection**
- **ML - algorithms to efficiently respond to potential network threats in real time**
- **“Big Data”:** acquisition, storage, analysis, transfer, visualization, querying, privacy
- **Human *in, on, out* of the loop?** (remote ctrl, semi-autonomous, autonomous). **“Meaningful” human control**
- **Cyber-intelligence?** Current algorithms not capable of human level reasoning. Presently employed to process & manage (sensor) data, monitor systems integrity, support vocal commands, navigate
- **Support C4ISTAR system: Command, Control, Communications, Computing, Information, Intelligence, Surveillance, Targeting Acquisition & Reconnaissance**

We already have weapons that can use AI to search, select and engage targets in specific situations

AI intrinsic vulnerabilities

- Growing pervasiveness gives rise to “adversarial AI”: exploiting machine learning models to misinterpret inputs into the system and behave in a way that’s favorable to the attacker
- Produce unexpected behavior: attackers create “adversarial examples” often appearing as normal inputs but instead meticulously optimized to break the model (instability & inaccurate predictions)
- Exploiting a particular behavior in AI internals (Neural Networks) unknown to developers
- Opacity of AI / Machine Learning / Deep Learning internals (NN level). Black box model (even designers cannot explain why an AI system reaches a specific result)
- “Poisoning attacks” during *supervised learning* phase (wrong, noisy, manipulated, non balanced data). Biases (intellectual, ethical)
- Slowly time drifting conditions in *unsupervised learning*
- Backdoors
- Various classes of vulnerabilities
- AI algorithms currently lack of interpretability, predictability, verifiability, reliability

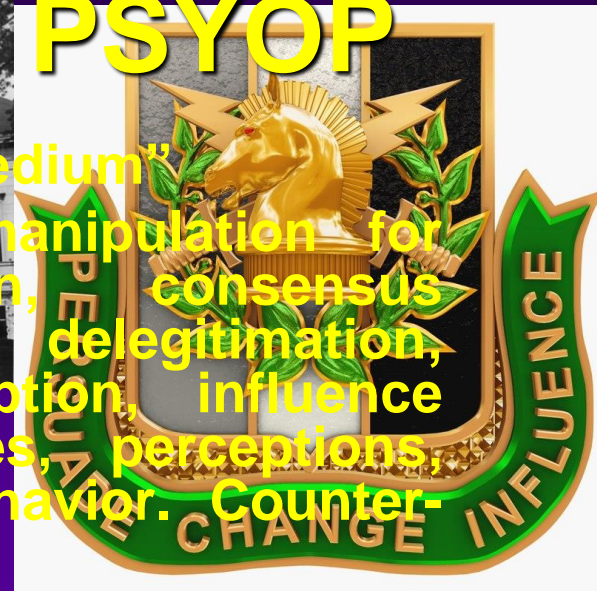
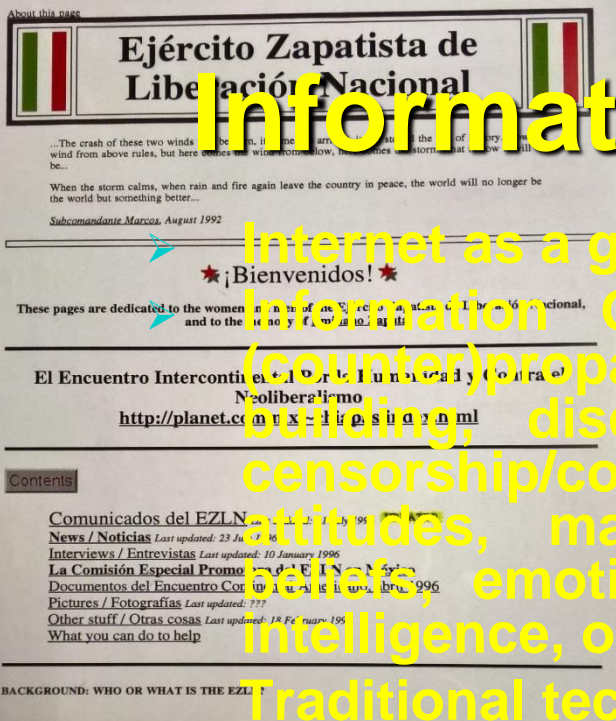
AI = Artificial Intelligence or Automating Ignorance??

- XAI (eXplainable Artificial Intelligence): evolution of AI such that results can be understood by humans
- Securing AI: 1)secure AI infrastructure 2)secure algorithms 3) secure training data 4)identify & manage external data dependencies

LAWS

- “Asilomar AI principles” developed by Future of Life Institute (FLI) during 2017 Asilomar conference (videos). Catalyze/support research & initiatives for safeguarding life and develop visions of the future, including positive ways for humanity to steer its own course considering new technologies and challenges
- “AI has already provided beneficial tools that are used every day by people around the world. Its continued development, guided by the following principles, will offer amazing opportunities to help and empower people in the decades and centuries ahead”
- 23 principles:
 - 5 on research domain (goals, funding, culture ...)
 - 13 on ethics and values (safety, responsibility, privacy, human control ...)
 - 5 on longer term issues (risks, common good ...)

March 2021: Lybia, first recorded case of an autonomous drone attack(?!) (STM Kargu-2). UN report: “The deployment to Lybia by Turkey is in non-compliance with par.9 of resolution 1970 (2011)”



Information Warfare e PSYOP

Internet as a global communication “medium”

Operations (IO): info manipulation for consensus building, discrimination, defamation, delegitimation, censorship/content filtering. Deception, influence attitudes, manipulate target's values, perceptions, beliefs, emotions, reasoning and behavior. Counter-intelligence, ops security

Traditional techniques (centuries old) on a new medium

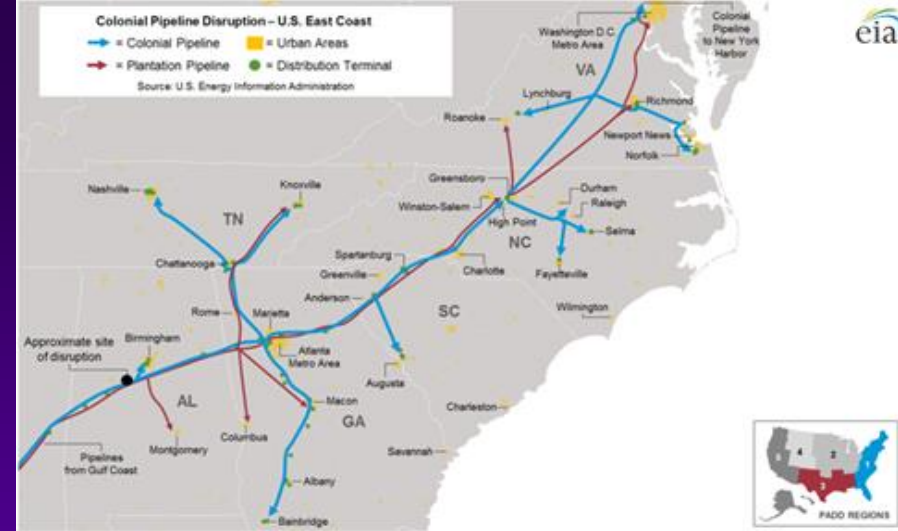
“Nihil est quod videtur” “..Cicero..”

- Real world examples: support to dissident groups, recruitment campaigns, use/manipulation of social media/networks. Wikileaks (2010, Assange), NSA Leaks (2013, Snowden), CIA Leaks (2017), EZLN ('90)
- Network is an ubiquitous surveillance environment
- Info war: primary political (strategic) value. “cyber influence” might contribute to political and social instability of a country. Blurring distinction between military and civilian domains

Ransomware

Colonial Pipeline May 7th 2021

- **Largest pipeline in US providing ~45% of the East Coast's fuel, including gasoline, diesel, home heating oil, jet fuel, & military supplies. Attack blocked oil production & commerce for several days, some outages/risk of shortages**
- **Linked to DarkSide group (RaaS paradigm); released a statement claiming that "our goal is to make money, and not creating problems for society"**
- **Access through remote desktop protocols (list of forbidden organizations to attack and eastern language op.sys.). Data exfiltration for double-extortion followed by encryption**
- **DarkSide targeted the business side rather than operational systems (intent was money-orientated rather than crashing down pipeline); CP "proactively took certain systems offline to contain the threat, which temporarily halted all pipeline operations", resuming (some) manual control**
- **ICS & CI high-value targets. Close monitoring by White House as disruption to the supply lines for potentially a full week, could lead to supply problems for consumers, aviation & military**
- **CP restart operations on May 12th after paying \$4.4 million (75 Bitcoins) ransom. On June 7th US DoJ announced recovery of \$2.3 million (63.7 Bitcoins) without disclosing details**



Cyber diplomacy



Going forward:

in 2018 established two parallel (hopefully converging & complementary) processes to discuss ICT security in 2019-2021

- **Open-Ended Working Group (OEWG) open to all Member States (chair Amb.Jürg Lauber CH). Report to GA in 2020. OEWG will hold an inter-sessional consultative meeting with industry, civil society, NGOs and academia (new wider approach)**
- **New GGE of 25 members (chaired Amb.Guilherme de Aguiar Patriota BR). Final report in 2021. The Chair will hold consultations with the wider membership in between sessions. Consultations with regional organizations (AU, EU, OAS, OSCE, ASEAN)**
- **OEWG should refer to shared conclusions of previous GGEs (2015 A70/174). OEWG represents by itself a sort of CBM
Final report (march 2021)**
- **“The right to privacy in the digital age”, A/RES/68/167 (2013)**
- **Proposed universal code of conduct for Information Security (2015)**
- **G7 Declaration on responsible state behavior in cyberspace (2017)**

OEWG final report (A/AC.290/2021/CRP.10 March 2021)

- **“Number of states are developing ICT capabilities for military purposes”**
- **Malicious use of ICTs by state and non-state actors (including terrorists & criminal groups). Some non-state actors with ICT capabilities previously only available to states**
- **Devastating security, economic, social, humanitarian consequences of malicious ICT activities on CI & CII (often owned by private sector. Cooperation needed)**
- **Lack of awareness & capacities to detect, defend against or respond to malicious ICT activities. No state is sheltered from threats**
- **Implement rules, norms & principles for responsible state behavior. Ensure integrity of supply chain, prevent proliferation of malicious ICT tools, reporting of vulnerabilities**
- **International law (UN Chart) applicable in ICT environment. Contribute to building consensus & common understandings within international community**
- **Confidence Building Measures (CBMs) at bilateral, regional, multilateral level for transparency, cooperative & stability measures contributing to prevent conflicts, avoid misperception/misunderstandings, reduction of tensions**
- **Establish communications, building bridges & initiating cooperation on shared objectives of mutual interest. National Points of Contact (PoCs)**
- **Capacity building (technical assistance, coordination CERTs, CSIRTs)**
- **New OEWG on ICT established for 2021-2025**

UN GGE 2019/21 final report (28 May 2021)

- **Existing and emerging threats: serious concerns about**
 - harmful ICT activity against critical infrastructure
 - increase in states' malicious use of ICT-enabled covert information campaigns to influence the processes, systems and overall stability of another state
 - malicious ICT activity aimed to exploit vulnerabilities
- **Norms, rules, and principles:**
 - reflect the expectations of the international community and set standards for responsible state behavior
 - additional norms could be developed over time, and, if appropriate, additional binding obligations could be elaborated in the future
 - group has also developed an additional understanding of the 13 voluntary GGE 2015 norms
- **International law**
 - reaffirmed applicability of international law & and UN Charter to the ICT environment
 - clarified that IHL applies only in situations of armed conflict
- **Confidence building measures**
 - cooperative (points of contact (PoC) and dialogue and consultations) & transparency measures (bilateral, sub-regional, regional, multilateral fora and informal consultations to clarify positions/share information)
- **International cooperation and assistance in ICT security and capacity-building**
 - support states in developing/implementing national ICT policies, strategies & programmes. Creating/enhancing capacity of CERTs/CSIRTs and their cooperation. Harmonization with OEWG report



2021 Cyber Stability Conference: Towards a More Secure Cyberspace

Preparatory conference for first meeting of new OEWG 2021-2025

➤ Existing and potential threats

How important will it be for a 5-year(!) process to monitor and account for the rapid evolution of the technological landscape?

To what extent should the OEWG also cover evolving threats from non-state actors?

➤ Rules, norms and principles for responsible state behavior

Voluntary and non-binding norms of responsible State behaviour can reduce risks to international peace, security & stability. Which norms are in need of further development?

How can industry and other non-state actors support OEWG in the norm development process?

➤ International law

Is there a role for non-State actors and/or for the International Law Commission?

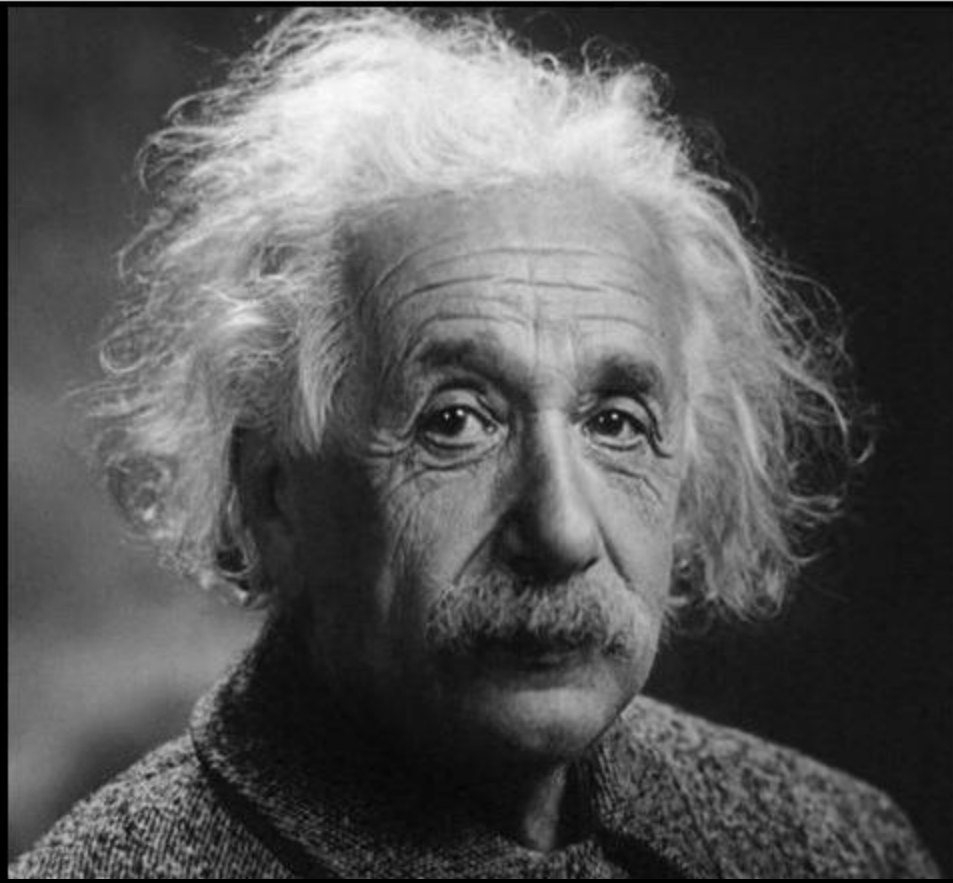
➤ Confidence-building measures

How can OEWG support States engaging in transparency measures, as with the sharing of relevant information and lessons learned?

How can the OEWG leverage initiatives led by non-state actors (civil society/NGOs, private sector & technical community) that could contribute to shared goals of transparency, information sharing & cooperation?

➤ Capacity building

How can OEWG support South–South, South–North, triangular & regionally focused cooperation in cyber capacity-building? What is missing from the discussion on models for cooperation towards cyber capacity-building?



Solution is not at the ICT technical level only

“We cannot solve problems by using the same kind of thinking we used when we created them” A.Einstein